The Genetic Code as an Admissibility Algebra: Translation Quotients, Rank-10 Physicochemical Projection, and the G-A-C-U Topological Policy

Douglas C. Youvan doug@youvan.com www.youvan.ai July 5, 2026

The standard genetic code is usually represented as a lookup table $\phi:\mathcal{C}\to\mathcal{R}$ from sixty-four codons to twenty amino acids and stop. This representation is biologically useful but algebraically nonminimal. We reformulate the code as an admissibility algebra over codon space. The primitive object is not the table alone, but the policy-object $\mathcal{G}$ = ($\mathcal{C}$, ϕ, D, Q, o), where $\mathcal{C} = \mathcal{A}^3$, ϕ is translated identity, D is the 64 × 12 nucleotide-position incidence matrix, Q assigns physicochemical values to translated labels, and o is an ordered nucleotide policy. Translation becomes a quotient $\mathcal{C}/\pi\phi$. Physicochemical structure becomes a rank-10 projection $PD = DD^{+}$. Synonymous degeneracy becomes stabilizer structure under translated identity. The ordered G-A-C-U cube becomes a topological policy whose 144 nearest-neighbor edges collapse to a 21-vertex multigraph with maximal synonymous self-loops. This unifies hydropathy, volume, pseudoinverse mapping, Hamming geometry, codon degeneracy, and G-A-C-U topology as different certified projections of one admissibility-regulated genetic-code object.

Keywords: genetic code, admissibility algebra, translation quotient, codon space, rank-10 codon matrix, exact pseudoinverse, nucleotide-position projection, hydropathy, residue volume, Hamming geometry, synonymous degeneracy, G-A-C-U, codon cube, topological policy, amino-acid multigraph.

1. Introduction

The standard genetic code is usually written as a map:

$\phi:\mathcal{C}\to\mathcal{R}$.

Here $\mathcal{C}$ is the set of sixty-four codons and $\mathcal{R}$ is the set of twenty amino acids plus stop X.

This is correct but not structurally minimal. The map ϕ is one policy on codon space: translated identity.

Admissibility Algebra permits a compressed reformulation.

The genetic code is a policy-object:

$\mathcal{G}$std = ($\mathcal{C}$, ϕ, D, Q, o).

$\mathcal{C} = \mathcal{A}^3$.

$\mathcal{A}$ = {G,A,C,U}.

ϕ:$\mathcal{C}\to\mathcal{R}$.

D ∈ $\{0,1\}^{64\times12}$.

Q:$\mathcal{R}\to\mathbb{R}^m$.

o = G-A-C-U.

The object $\mathcal{G}$std has several induced structures:

translation quotient,

rank-10 additive projection,

Hamming kernel,

ordered cube,

translated multigraph,

synonymous self-loop extremum.

These are not separate representations. They are policy-relative projections of one object.

2. Primitive Codon Space

Definition 2.1 Codon alphabet

Let:

$\mathcal{A}$ = {G,A,C,U}.

Definition 2.2 Codon space

Let:

$\mathcal{C} = \mathcal{A}^3$.

Thus:

$|\mathcal{C}| = 4^3 = 64$.

A codon is $c = c_1c_2c_3$ with $c_i \in \mathcal{A}$.

Definition 2.3 Translated label space

Let:

$\mathcal{R} = AA_{20} \cup \{X\}$.

Here AA_{20} denotes the twenty standard amino acids, and X denotes translation stop.

Definition 2.4 Genetic-code map

The standard genetic code is:

$\phi:\mathcal{C}\to\mathcal{R}$.

This map induces translated identity.

3. Translation Policy

Definition 3.1 Translated identity policy

For c,d ∈ $\mathcal{C}$:

c ≡ϕ d iff ϕ(c) = ϕ(d).

This is the first admissibility policy of the genetic code.

Definition 3.2 Translation quotient

The translated quotient is:

$\mathcal{C}/\pi\phi \cong \mathcal{R}$.

This quotient collapses synonymous codons and identifies the three stop codons as X.

Proposition 3.3

The ordinary genetic-code table is the quotient representation of codon space under translated identity.

Proof

The table records only ϕ(c). Two codons with identical translated labels are indistinguishable in the quotient. Therefore the lookup table is $\mathcal{C}/\pi\phi$ together with representative codon assignments.

Corollary 3.4

Synonymous degeneracy is stabilizer structure under πϕ.

4. Nucleotide-Position Presentation

Definition 4.1 Incidence matrix

Let D ∈ $\{0,1\}^{64\times12}$ be the complete codon incidence matrix.

Columns are indexed by position-nucleotide pairs:

(p,s) ∈ {1,2,3} × $\mathcal{A}$.

Rows are indexed by codons c ∈ $\mathcal{C}$.

D c,(p,s) = 1 iff c_p = s.

D c,(p,s) = 0 otherwise.

Each row contains exactly three ones.

Definition 4.2 Additive nucleotide-position policy

πD is the policy under which codon properties are considered identical up to their projection onto independent nucleotide-position contributions.

This policy ignores higher-order interaction terms among positions.

Theorem 4.3 Rank

rank(D) = 10.

Proof sketch

Each positional block has four columns whose sum is the all-ones vector. The three block-sum vectors are equal. These two independent dependencies reduce the apparent twelve-dimensional coordinate system to dimension ten.

Definition 4.4 Exact pseudoinverse

For the complete 64-codon matrix:

$D^+ = 1/16\ D^T - 1/96\ J$.

Here J is the 12 × 64 all-ones matrix.

Corollary 4.5

The complete codon incidence system has an exact analytical inverse policy in the Moore–Penrose sense.

5. Physicochemical Property Projection

Definition 5.1 Amino-acid property map

Let:

$Q:\mathcal{R}\rightarrow\mathbb{R}^m$.

Typical columns include hydropathy and residue volume. Stop X may be assigned neutral padding values when a complete 64-codon property field is required.

Definition 5.2 Codon property field

Let $Y \in \mathbb{R}^{64 \times m}$ be the codon-level property field induced by ϕ and Q.

$Y(c) = Q(\phi(c))$.

Equivalently:

$Y = \Phi Q$,

where Φ is the 64×21 codon-to-label assignment matrix.

Definition 5.3 Nucleotide-position coefficients

The additive nucleotide-position coefficient matrix is:

$B = D^{+}Y$.

Definition 5.4 Reconstructed additive field

The rank-10 additive reconstruction is:

$\hat{Y} = DB$.

Thus:

$\hat{Y} = DD^{+}Y$.

Let:

$PD = DD^{+}$.

Then:

$\hat{Y} = PDY$.

Proposition 5.5

Physicochemical information decomposes as:

$Y = PDY + (I - PD)Y$.

The first term is nucleotide-position recoverable.

The second term is residual interaction structure.

Corollary 5.6

Hydropathy and residue volume are not separate analyses. They are columns of the same property-field projection.

6. Hamming Kernel

Definition 6.1 Codon Hamming distance

For c,d ∈ $\mathcal{C}$:

$d_H(c,d) = |\{p : c_p \neq d_p\}|$.

Theorem 6.2 Projection kernel

For the complete codon incidence projection:

$PD(c,d) = 5/32 - d_H(c,d)/16$.

Thus:

$d_H = 0$ gives 5/32.

$d_H = 1$ gives 3/32.

$d_H = 2$ gives 1/32.

$d_H = 3$ gives −1/32.

Interpretation

The additive nucleotide-position projection is order-independent. It depends only on symbolic agreement among codon positions.

Corollary 6.3

The rank-10 additive policy πD is a Hamming-distance policy.

7. Second-Position Physicochemical Policy

Definition 7.1 Middle-base partition

Define:

$\mathcal{C}s = \{c \in \mathcal{C} : c_2 = s\}$.

for s ∈ $\mathcal{A}$.

The four blocks are:

$\mathcal{C}G$, $\mathcal{C}A$, $\mathcal{C}C$, $\mathcal{C}U$.

Proposition 7.2

The second codon position is the dominant coarse physicochemical classifier.

Policy interpretation

π_2 partitions codon space by middle nucleotide before translation.

Under π_2:

NUN is hydrophobic-enriched.

NAN is hydrophilic-enriched.

NCN is small-residue-enriched.

NGN is heterogeneous.

Definition 7.3

The second-position hydropathy certificate is the coefficient contrast:

$\beta_{2,U} - \beta_{2,A}$.

In the historical sense-codon model, this contrast is large and positive.

Corollary 7.4

Hydropathy is a translated property whose strongest additive coordinate is a nucleotide-position invariant.

8. Ordered Topological Policy

The incidence matrix D is order-independent. It does not know whether G is adjacent to A, C, or U.

Topology requires an additional policy.

Definition 8.1 Nucleotide order

Let o be a linear order on $\mathcal{A}$.

For the present code:

o = G-A-C-U.

The reversed order is:

o^r = U-C-A-G.

Definition 8.2 Ordered codon cube

Given o, define the ordered cube:

$K(o) = [4] \times [4] \times [4]$.

Each axis represents one codon position, and each coordinate value is assigned to a nucleotide by o.

Definition 8.3 Ordered nearest-neighbor edge set

Eo contains unordered codon pairs {c,d} such that c and d differ in exactly one position and their differing nucleotides are adjacent in o.

For a $4 \times 4 \times 4$ cube:

$|Eo| = 3(4-1)4^2 = 144$.

Definition 8.4 Topological policy

πo is the policy under which admissible codon substitutions are nearest-neighbor moves in K(o).

9. Collapse to Translated Multigraph

Definition 9.1 Translated multigraph

Let M(o,ϕ) be the multigraph obtained by collapsing each codon c to ϕ(c), while retaining all edges in Eo.

Thus:

$V(M) = \mathcal{R}$.

$|V(M)| = 21$.

Edges are inherited from Eo.

Definition 9.2 Synonymous self-loop

An edge $e = \{c,d\} \in Eo$ is synonymous iff:

$\phi(c) = \phi(d)$.

It becomes a self-loop in M(o,ϕ).

Definition 9.3 Self-loop indicator

$I\phi(e) = 1$ if $\phi(c) = \phi(d)$.

$I\phi(e) = 0$ otherwise.

Definition 9.4 Loop count

L(o) = ∑e∈Eo Iϕ(e).

Theorem 9.5 G-A-C-U extremum

Among all twenty-four isotropic nucleotide orders:

maxo∈S_4 L(o) = 44.

The maximizing oriented orders are:

G-A-C-U

and

U-C-A-G.

Corollary 9.6

G-A-C-U is an extremal topological policy for synonymous compactness.

10. Unified Admissibility Representation

The genetic code is now represented by one object:

$\mathcal{G}$std = ($\mathcal{C}$, ϕ, D, Q, oGACU).

Its induced policies are:

πϕ = translated identity.

πD = rank-10 additive nucleotide-position projection.

πH = order-independent Hamming projection.

π_2 = middle-base physicochemical partition.

πo = ordered nearest-neighbor topology.

πM = translated multigraph collapse.

Each policy produces a different certified representation.

Definition 10.1 Genetic-code admissibility object

$\mathcal{G}$std is the minimal object whose derived structures include:

$\mathcal{C}$/πϕ,

PD = DD^+,

$B = D^{+}Y$,

$PD(c,d) = 5/32 - dH(c,d)/16$,

K(G-A-C-U),

M(G-A-C-U,ϕ),

L(G-A-C-U) = 44.

Theorem 10.2 Compression theorem

The table, matrix, property projection, Hamming geometry, ordered cube, and amino-acid multigraph are all projections or quotients of $\mathcal{G}$std under distinct admissibility policies.

Proof sketch

The table is induced by ϕ. The matrix is induced by D. Physicochemical fields are induced by $Q \circ \phi$. Additive reconstruction is induced by D^{+}. Hamming geometry is induced by PD. The cube is induced by o. The multigraph is induced by ϕ applied to Eo. Therefore all structures arise functorially from $\mathcal{G}$std.

11. Biological Interpretation

The genetic code is usually described as degenerate. Under Admissibility Algebra, degeneracy becomes policy-stable identity.

A synonymous codon set is an orbit under πϕ.

A fourfold codon family is a local stabilizer of translated identity under third-position substitution.

A sixfold family is a union of translated orbits across more than one codon block.

A stop signal is a translated label X with its own quotient class.

The hydropathy signal is not an external annotation. It is a property field over the quotient lifted back to codon space.

The G-A-C-U maximum is not merely a graphical fact. It is an extremal identity-preservation result under πo.

Thus the code is simultaneously:

a translation quotient,

a physicochemical projection,

a symbolic Hamming geometry,

and an ordered topological stabilizer system.

12. Conclusions

The admissibility formulation simplifies the genetic code by removing representational fragmentation.

The conventional table is $\mathcal{C}/\pi\phi$.

The hydropathy and volume analyses are columns of Y = ΦQ.

The nucleotide-position coefficients are B = $D^{+}Y$.

The recoverable additive component is PDY.

The projection kernel is determined by Hamming distance.

The ordered topology is generated only after choosing o.

The G-A-C-U order is a maximal synonymous-compactness policy.

The complete compressed statement is:

$\mathcal{G}$std = ($\mathcal{A}^3$, ϕ, D, Q, G-A-C-U).

with:

rank(D) = 10.

D^{+} = 1/16 D^{T} − 1/96 J.

PD = DD^{+}.

PD(c,d) = 5/32 − dH(c,d)/16.

Y = ΦQ.

B = $D^{+}Y$.

$\hat{Y}$ = PDY.

L(G-A-C-U) = L(U-C-A-G) = 44.

This is the simplified representation.

The genetic code is not merely a lookup table with hidden structure.

It is an admissibility algebra whose biological function, physicochemical organization, symbolic geometry, and topological compactness are certified projections of one finite codon object.

13. References

1. Youvan, D. C. The Genetic Code as Physicochemical and Topological Geometry: Nucleotide-Determined Hydropathy, Exact Pseudoinverse Mapping, and the G-A-C-U Maximum. ResearchGate preprint, 2026.
2. Yang, M. M., Coleman, W. J., and Youvan, D. C. Genetic coding algorithms for engineering membrane proteins. In Reaction Centers of Photosynthetic Bacteria. Springer Series in Biophysics, Vol. 6. Springer-Verlag, 1990.
3. Kyte, J., and Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology, 157, 105–132, 1982.
4. Zamyatnin, A. A. Protein volume in solution. Progress in Biophysics and Molecular Biology, 24, 107–123, 1972.
5. Moore, E. H. On the reciprocal of the general algebraic matrix. Bulletin of the American Mathematical Society, 26, 394–395, 1920.
6. Penrose, R. A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society, 51, 406–413, 1955.
7. Mac Lane, S. Categories for the Working Mathematician. 2nd ed. Springer, 1998.
8. HoTT Book Authors. Homotopy Type Theory: Univalent Foundations of Mathematics. Institute for Advanced Study, 2013.
9. Youvan, D. C. Admissibility Algebra: A Policy-Relative Foundation for Identity, Invariance, and Certified Transformation. ResearchGate preprint, 2026.